

OPTIMIZATION OF SELECTED REMOTE SENSING ALGORITHMS FOR EMBEDDED NVIDIA KEPLER GPU ARCHITECTURE

Lubomír Říha

IT4Innovations National Supercomputing Center, VŠB-Technical University of Ostrava, Ostrava, Czech Republic

Jacqueline Le Moigne

NASA Goddard Space Flight Center, Software Engineering Division, Greenbelt, MD, USA

Tarek El-Ghazawi

High-Performance Computing Laboratory, The George Washington University, Ashburn, VA, USA

This paper evaluates the suitability of new embedded Graphic Processing Units with 192 single precision cores (GPU) in the Nvidia’s Tegra K1 (K1) System-on-Chip (SoC) with typical Typical Design Power (TDP) under 7W [1] for onboard processing. The performance of this SoC is compared to two modern High Performance Computing (HPC) architectures:

- (1) General purpose multi-core CPU (8-core Sandy Bridge E5-2470, 2.3GHz, TDP 95W [2])
- (2) GPU accelerator (Nvidia Tesla K20 (K20), TDP 225W [3]).

For this study, we selected two algorithms:

Wavelet Spectral Dimension Reduction of Hyperspectral Imagery

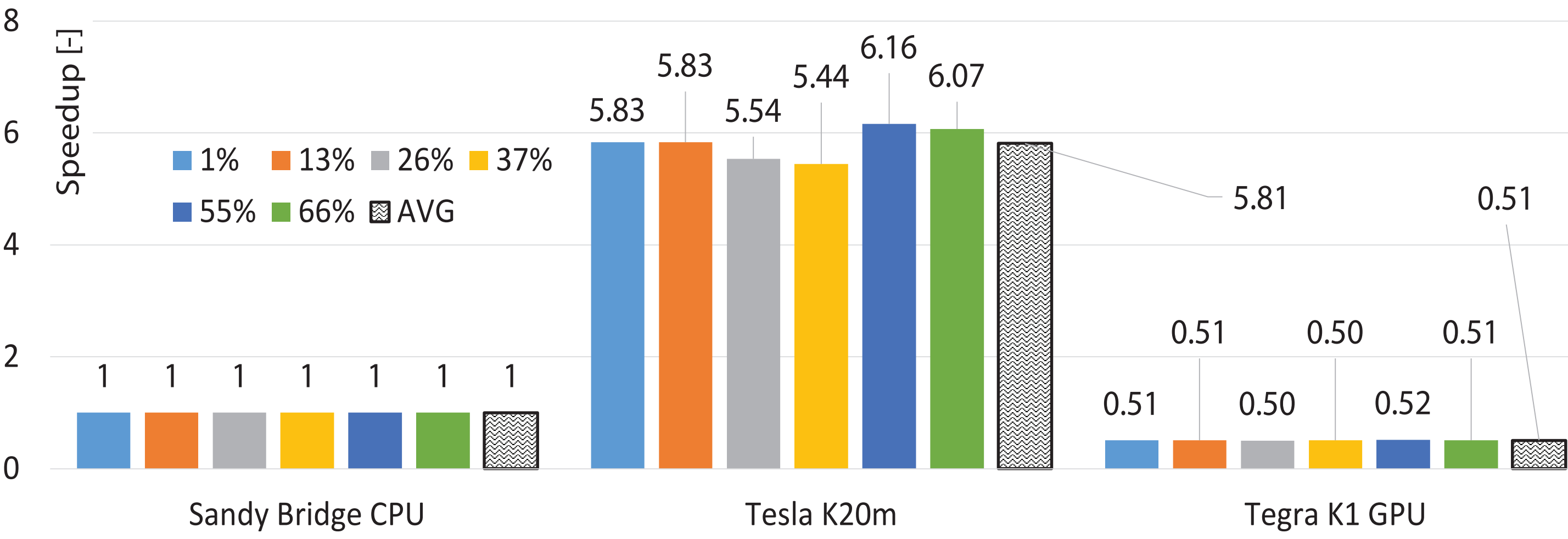
The principle of this method is to apply a discrete wavelet transform to hyperspectral data in the spectral domain and at each pixel location. The optimal level of wavelet decomposition is computed adaptively for each pixel. See [4] for more details.

Automated Cloud-Cover Assessment (ACCA) Algorithm

The ACCA algorithm determines and rates the overall cloud cover of an image through 2 steps: Pass-One isolates clouds from non clouds by utilizing eight threshold-based filters, then Pass-Two resolves the detection ambiguities from Pass-One by computing global statistics over the image. See [5] for more details.

This paper shows that the performance achieved using this new SoC designed for battery powered devices is comparable to HPC hardware with significantly higher power consumption.

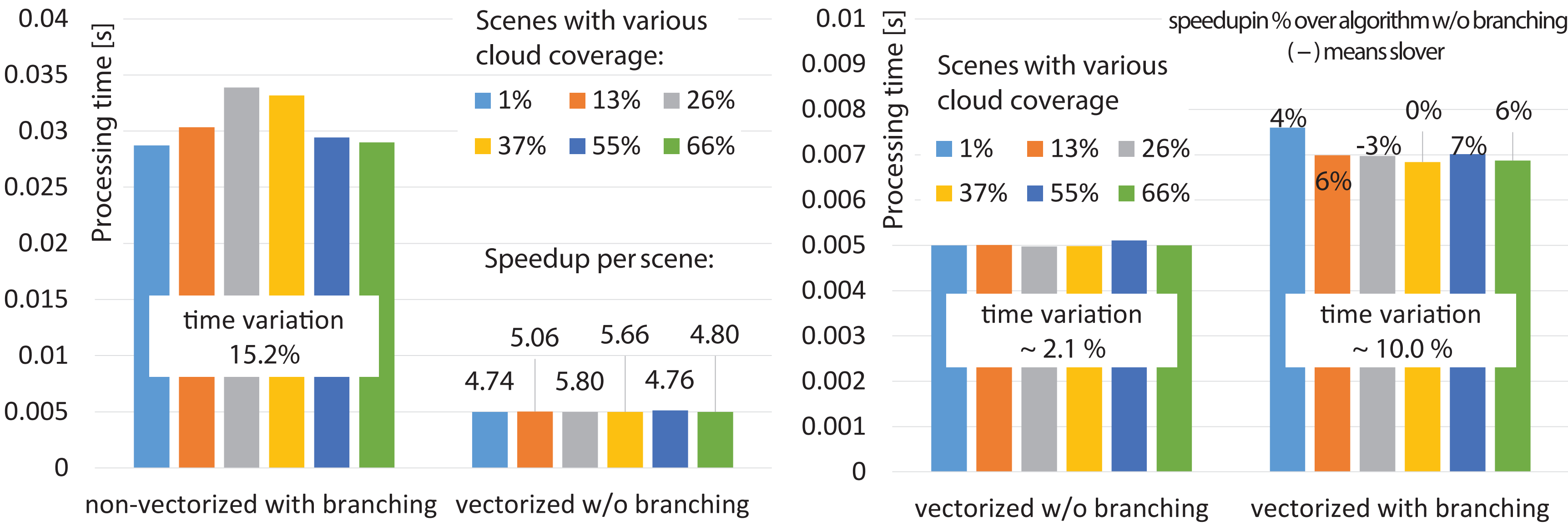
In order to gain optimal performance we had to redesign the original algorithms to support SIMD processing. Tegra K1 achieved (1) 51% for ACCA algorithm and (2) 20% for the dimension reduction algorithm, as compared to the performance of the high-end 8-core server Intel Xeon CPU. Both algorithms use only a GPU part of the SoC, leaving the 4+1 ARM Cortex A15 general-purpose cores available for other tasks.



Chip-to-chip performance comparison of the vectorized ACCA algorithm without branching for image size 2048x2048 pixels.

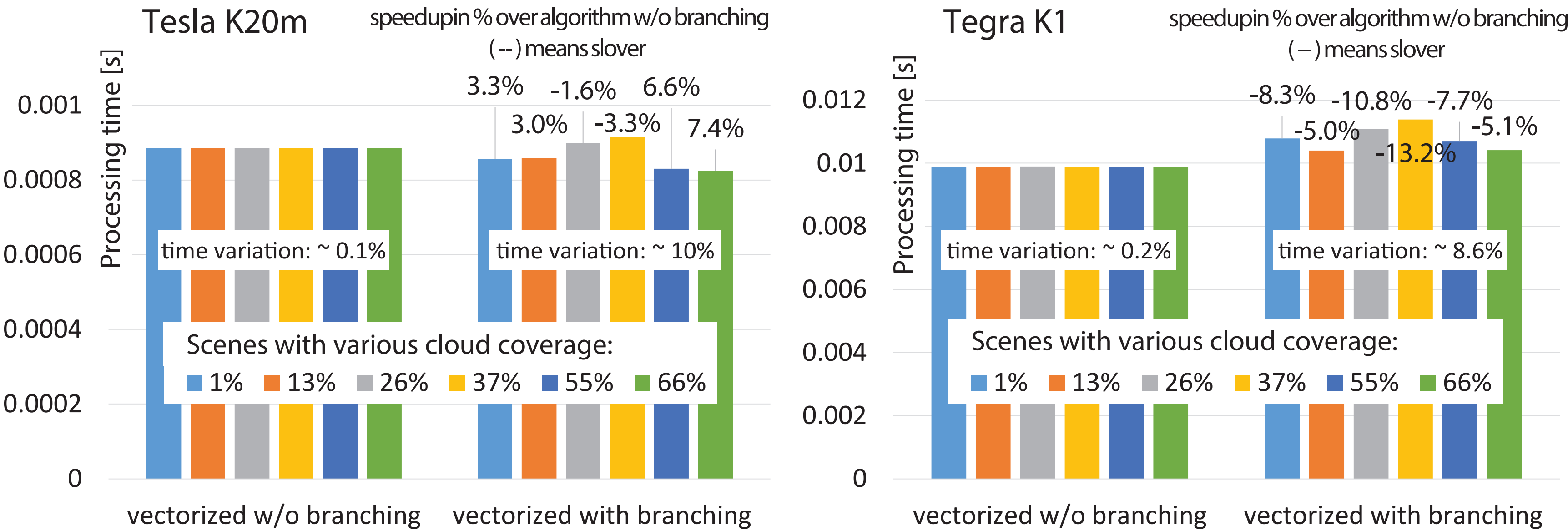
	Nvidia Tegra K1 (GPU part)	8-core Intel Sandy Bridge E5-2470	Nvidia Tesla K20 GPU
Architecture type	embedded SoC with Kepler GPU	general purpose CPU for HPC	GPU accelerator for HPC
Frequency	0.852 GHz – GPU part	2.3GHz	0.706GHz
Number of Cores	192 SP scalar cores – GPU part	64 SP / 32 DP cores (8 SIMD cores)	2496 SP / 832 DP scalar cores
On-Chip Caches	64 KB L1 per 192 SP cores 128KB L2 per chip	32+32KB L1, 256 KB L2 per SIMD core 20 MB L3 per chip	64 KB L1 per 192 SP cores; 1536KB L2 per chip
SIMD width	32 for both SP and DP	8 for SP and 4 for DP	32 for both SP and DP
Peak Performance	327 SP / 13 DP GFLOPS	147 SP / 74 DP GFLOPS	3524SP/1160DP GFLOPS
Mem. Size; Bandwidth	2GB at Jetson TK1; 14.9 GBPS	up to 384 GB; 38.4 GBPS	5GB; 208 GBPS
TDP	7W (SoC + DRSM only)	95W (CPU only)	225W (accelerator only)

Main parameters of the selected hardware architectures

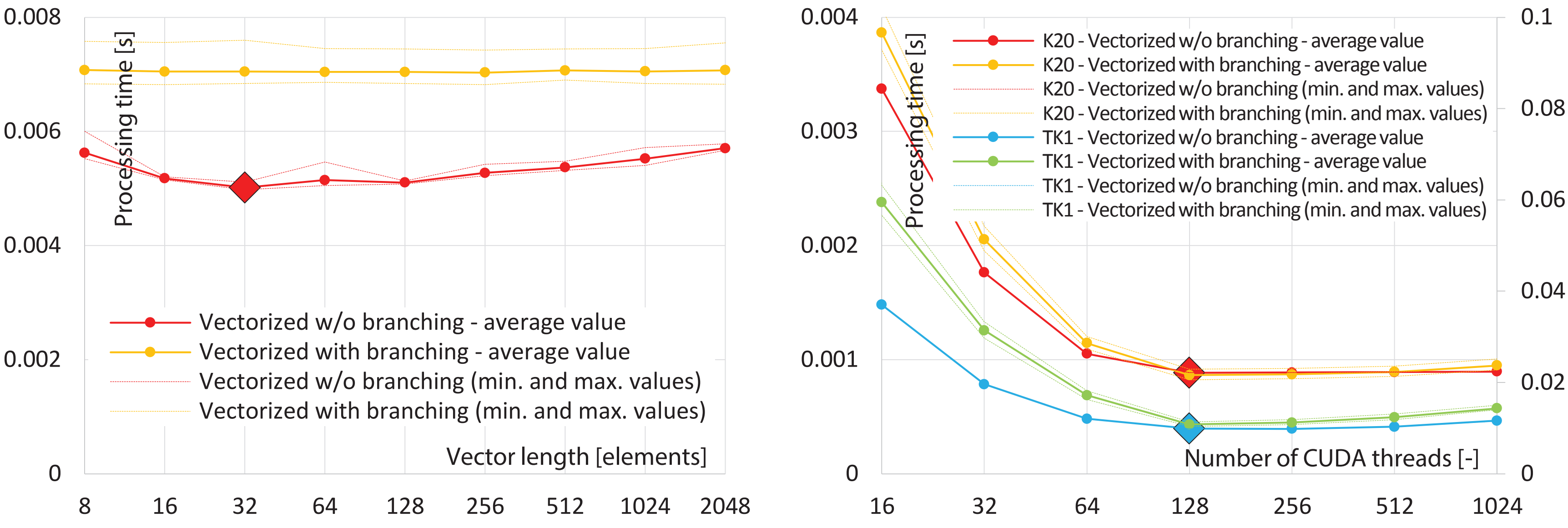


Speedup achieved by vectorization for CPU is between 4.7 and 5.8. The processing time variation of original algorithm for different input data is 15.2%. The values above the bars show the speedup for different scenes with various cloud coverage from 1% to 66%.

Processing time variation based on input data with various cloud coverage 1%, 13%, 26%, 37%, 55% and 66% for CPU. The values above the bars describe the difference in processing time: negative values means slower than vectorized w/o branching algorithm.



Processing time variation based on input data with various cloud coverage 1%, 13%, 26%, 37%, 55% and 66% for Tesla K20 and Tegra K1 GPUs. The values above the bars describe the difference in processing time: negative values mean slower than vectorized no-branching algorithm. Image size is 2048x2048 pixels. .



Optimal vector length for the ACCA algorithm running on CPU is 32. Optimal number of threads per block for ACCA on Tesla K20 and Tegra K1 is 128.

Architecture	Spectral Bands per pixel [-]	Processing time [s]	Performance [Mpix per second]	Speedup over CPU [-]
8-core CPU	128	0.0303	3.14	1
	256	0.0394	2.42	1
	512	0.0627	1.52	1
Nvidia Tesla K20	128	0.0118	8.08	2.57
	256	0.0177	5.39	2.23
	512	0.0411	2.32	1.53
Nvidia Tegra K1	128	0.6133	0.61	0.20
	256	0.4225	0.42	0.19
	512	0.1793	0.18	0.17

Chip-to-chip performance comparison of the Wavelet Spectral Dimension Reduction algorithm for image size 100,000 pixels.

[1] Nvidia, “NVIDIA Tegra K1: A New Era in Mobile Computing”, http://www.nvidia.com/content/PDF/tegra_white_papers/Tegra-K1-whitepaper-v1.0.pdf, 2014.

[2] Intel, “Intel® Xeon® Processor E5-2400 Product Family”, <http://www.intel.com/content/dam/www/public/us/en/documents/datasheets/xeon-e5-2400-vol-1-datasheet.pdf>, May 2012.

[3] Nvidia, “NVIDIA’s Next Generation CUDATM Compute Architecture: Kepler TM GK110”, <http://www.nvidia.com/content/PDF/kepler/NVIDIA-kepler-GK110-Architecture-Whitepaper.pdf>, 2012.

[4] S. Kaewpijit, J. Le Moigne, T. El-Ghazawi, "Automatic reduction of hyperspectral imagery using wavelet spectral analysis," Geoscience and Remote Sensing, IEEE Transactions on , vol.41, no.4, pp.863,871, 2003.

[5] R. R. Irish, et al. "Characterization of the Landsat-7 ETM+ automated cloud-cover assessment (ACCA) algorithm." Photogrammetric Engineering & Remote Sensing 72.10 (2006): 1179-1188.

[6] Nvidia, “NVIDIA Jetson TK1 Development Kit”, http://developer.download.nvidia.com/embedded/jetson/TK1/docs/Jetson_platform_brief_May2014.pdf, 2014.